

<https://helda.helsinki.fi>

Transparency in the Digital Environment

Koivisto, Ida

2021-04-03

Koivisto , I 2021 , ' Transparency in the Digital Environment ' , Critical Analysis of Law , vol. 8
, no. 1 , pp. 1-8 .

<http://hdl.handle.net/10138/328816>

cc_by
publishedVersion

Downloaded from Helda, University of Helsinki institutional repository.

This is an electronic reprint of the original article.

This reprint may differ from the original in pagination and typographic detail.

Please cite the original version.

Transparency in the Digital Environment

Ida Koivisto*

Abstract

In this Editor's Introduction to the CAL Special Issue on "Transparency in the Digital Environment," I discuss some of its overarching themes and underlying concerns. I also present the eight articles that constitute this issue: What kinds of manifestations can transparency as an idea have in the digital environment? Transparency has proven to be a normatively attractive concept in ethical and legal debates of automated decision-making, data protection, platform governance, artificial intelligence, and so on. In many of them, the recurring worry is this: Who is accountable for the functioning of algorithmic systems? How can we know how these systems work? By way of an answer, transparency is often considered the solution. However, although we worry similarly about accountability in both analog and digital contexts, the latter poses technological, regulatory, and ethical challenges, which may not always find a perfect counterpart in the analog environment. I show how the articles in this special issue represent three different approaches to transparency: reformative, performative, and conceptual. I conclude that we should continue critical discussion on the big words that dominate the current debates of legitimacy in the digital environment.

I. Introduction: Transparency Goes Digital

Some years ago, I saw an art exhibition by Taryn Simon, an American photographer. Among other works of art, "The Innocents" (2003), a series of photographs, was on display. Each photograph depicted a reconstruction of a crime scene, showing the place where the crime had taken place and the person who had supposedly committed it. The photographs came with captions that explained this idea. Otherwise, the photos would have merely shown different people in different environments: a man hiding under a mattress, a man in a bar, a man and a woman against a brick wall. The catch was that the people in the pictures were convicted based on eyewitness evidence. Later, DNA testing—a technology that was not available at the time of the conviction—proved them innocent. In some cases, this happened after decades of confinement.

I was perplexed, and it took me a while to understand why. I realized that the pictures constructed a disturbing half-truth; they showed wrongly convicted people amidst actual crime scenes after a considerable amount of time had passed. Additionally, the text made a significant difference. Without the captions, I would not have suspected that the person was in the middle of a crime scene. If I had not known that a crime had occurred, I

* Associate Professor of Public Law, University of Helsinki. My warmest thanks to Markus Dubber and Simon Stern for inviting me to guest-edit this special issue and to all the contributors. I also thank Trond Ove Tøllefsen and Riikka Koulu for their helpful comments on the earlier versions of the text and Linda Sydänmaanlakka for research assistance.

would have looked at the pictures differently. My gaze would be perhaps less inquisitive, less judging. This was not an accident: the artist had carefully planned the inevitable confusion that the viewer would experience. In an ingenious way, the artwork both made visible the institutionalized truth-value of seeing for oneself and at the same time questioned it. This ambivalence resided in the images, in their seeming immediacy, combined with the historical narrative in the captions.

As Simon's art illustrates, seeing for oneself is culturally considered the privileged way of knowing. Otherwise, we would not rely on eyewitness evidence. The cultural—perhaps even anthropological—value of seeing is so powerful that it can be questioned only with technology that goes beyond the human senses, such as DNA testing. As shown in "The Innocents," seeing and knowing may contradict each other. However, this contradiction easily remains a footnote in ocularcentric cultures: the power of seeing creates the default condition. As a token of this, *transparency*—one of the most powerful values of our time—is specifically premised on the power of seeing as a guarantee of knowing. It promises a clear view of reality. This simple promise has enabled it to become a norm: if we *can* see clearly, we *should* see clearly.

Over the last few decades, however, transparency requirements have grown exponentially. Transparency is no longer a sole matter of democratic theory and public law. Initially in those contexts, it empowered citizens to hold government accountable. Citizens had the right to know how power in a state was exercised. Recently, transparency has become an independent discourse of legitimacy. The expectation of transparency no longer seems to require particular justifications; the person who claims otherwise has the burden of proof. Transparency seeks to demystify and to simplify obscure and complex issues, to make the world rational and conceivable, yet it is a well-nigh mythical concept itself.¹ As Michael Schudson argues:

The growing approval for transparency or the cultural expectation of a right to know, and institutional reforms to embrace it, blend together a variety of values. The blend favors the open over the hidden, confidential or secret; expressiveness over politeness; the grassroots, egalitarian, or democratic over the elitist; the authority of personal experience over the authority of professional expertise; the natural over the adulterated; the accessible and plain-spoken over the recondite; and in general, freedom over restraint. These contrasts give favor to a utopian impulse, to the romantic over the classical in style, to a democratic faith in human beings.²

This cultural expectation of transparency has proven strong and resilient. It is no longer tied to any particular environment or context; it does not shy away from the contexts of technology. In fact, although transparency builds on the idea of immediate observability and firsthand knowledge, it is paradoxically necessarily a matter of mediation and technology—and mediating technology. Transparency's metaphorical promise cannot avoid

¹ Jan Teurlings & Markus Stauff, Introduction: The Transparency Issue, 14 *Cultural Stud. & Critical Methodologies* 3 (2014).

² Michael Schudson, *The Rise of the Right to Know: Politics and the Culture of Transparency, 1945-75*, at 21 (2015).

forming a tense relationship with mediation: the more the object of transparency becomes mediated, the less immediacy there can be.

Even a window can be regarded as a mediating technology. Mediation cannot but be a part of the representation of the thing mediated, and as such, part of the message delivered: we do not only see through a window, we see also the window itself. Even a window can distort the view, refract the light that passes through it, or simply break. In fact, we notice a broken window more easily than an intact one. Possibly for this very reason, transparency is called for: datafication shows clearly that something is broken, but we struggle to understand what that something is.³

Perhaps due to transparency's inherent dependency on mediation and technology, it is hardly surprising that transparency has found its way into current discourses on digitalization and algorithmization. It has proven to be normatively attractive in ethical and legal debates of automated decision-making, data protection, platform governance, and artificial intelligence, to mention only a few. The recurring worry is this: Who is accountable for the functioning of algorithmic systems? Although we worry about accountability in both analog and digital contexts in similar ways, the latter poses technological, regulatory, and ethical challenges, which may not always find a perfect analogy in the analog environment.

Two books especially map out the field of problems we encounter in the digital environment: Frank Pasquale's *The Black Box Society* (2015) and Shoshana Zuboff's *The Age of Surveillance Capitalism* (2019). These books have pinpointed some of the most pressing problems of the datafied society. The main concern in both is a capitalist logic that in the digital environment utilizes our nonchalant Internet behavior. Because we easily mistake the Internet as an anonymous place, the capitalist logic may infiltrate into our private realm and commodify our personal data. Often, however, we remain ignorant of these algorithmic operations. Accordingly, many of the digital transparency discussions build around the black box problem, a condition in which we do not know how inputs translate into outputs in algorithmic systems. If we could only open the boxes and see their inner workings! The black box has become an emblem of algorithmic ignorance.

The phenomena of the black box society and surveillance capitalism are specifically characteristic of the digitalized and datafied environment; a quite similar logic is absent from the analog world. Although transparency has traditionally been a value of public law and has been directed toward the state, its migration to ultimately private—though highly unequal—affairs may be a sign of new power configurations; transparency and power have always been closely connected. However, transparency is an intellectual invention of the analog world, whether it be tied to technology and mediation. Therefore, although transparency is so often called for, its potency to solve the problems of the datafied society is questionable, as is argued in this special issue.

³ I thank Riikka Koulu for this observation.

II. This Special Issue

This special issue delves into the ways in which the cultural expectation of transparency has found its way into the digital environment. The issue consists of eight articles by a multi-disciplinary group of scholars, mostly legal scholars. Some of them come to the discussion from a transparency studies perspective, others from a law and digitalization viewpoint. Although the contributors represent different scholarly traditions, each of them shows remarkable sensitivity to power.

In each article, the promise and premise of transparency in the digital environment are in focus. The underlying questions are thorny: Can we trust our eyes as a means of knowledge? How is digital reality represented to us through transparency practices? What difference does technology make? To what extent is transparency political? Is more information the solution to the problems of the black box society? How can we tackle the problems we identify? The digital environment cannot but effect how transparency works: visibility in the digital environment is governed not only through policy filters, but also through technological filters.⁴

Oana Brindusa Albu and Hans Krause Hansen focus on the idea of a face and its power implications in the digital environment. The authors approach the theme from three partially overlapping angles: datafied transparency, algorithmic governmentalities and biometric surveillance. Mostly, the authors construct their analysis based on assemblage theory, which investigates the historically contingent relations and linkages between humans and their material surroundings, including technologies. The authors use facial recognition as a practical example in which these three different theoretical insights coalesce. After the theoretical part, the authors go through different legislative examples and ask: How can facial recognition be legally tackled? The authors conclude that datafied transparency is inseparable from algorithmic governmentalities. This leads to algorithmic anxiety, whereby subjects cannot escape the surveilling gaze. Perhaps “algoactivism” is the solution?

Mark Fenster discusses the ways in which transparency needs to follow a particular legal script to be conceived of as officially reliable. To that end, social media platforms play a crucial role. This is illustrated with a unique case study: Donald Trump’s tweets about his health, more specifically about his COVID-19 infection. First, Fenster goes through the U.S. federal legislation on presidential health and shows that it leaves much leeway for obfuscation and cover-ups. After that, Fenster demonstrates the problematic nature of the narrative that Trump concocts through his tweets. Fenster argues that by bypassing the official transparency channels of the state, Trump creates an illusion of immediacy and honesty, which however proves to be irresponsible and misleading. Fenster concludes that Trump used his illness to create drama and gain attention on social media. In this way, Trump’s actions reveal many troubling features of transparency as an ideal.

⁴ Susan J. Drucker & Gary Gumpert, *Through the Looking Glass: Illusions of Transparency and the Cult of Information*, 26 J. Mgmt. Dev. 493 (2007).

Mateusz Grochowski, Agnieszka Jablonowska, Francesca Lagioia, and Giovanni Sartor address the values of algorithmic transparency and explainability in European Union consumer law. The authors go through different risks that a consumer encounters on the digital market: mass surveillance, manipulation and exploitation. How does the opacity of algorithmic influencing and decision-making affect the consumers' standing? The authors argue that the way in which transparency is understood in European consumer law and data protection regulation emphasizes certain aspects of transparency over others. Instead of providing understandability to experts (esoteric transparency), EU law focuses on furnishing adequate information to laypeople (exoteric transparency). The authors discuss the effect of this emphasis and conclude that the regulation of algorithms is still an immature element of EU law. It is mostly premised on information asymmetry between the consumer (who knows little) and the seller (who knows much) in the market, and the effort to rectify that asymmetry. However, this starting point has proven insufficient in the digital market, as the consumer's ability to make rational and informed choices—regardless of the quantity of the provided information—has been questioned in recent research. The digital environment thus poses particular challenges to EU law's fundamental premises.

Ida Koivisto in her article concentrates on transparency and its relation to digital reality production. In the wake of digitalization and datafication, we are increasingly oblivious to how digital realities are produced: To what extent can we rely on their truthfulness? Is transparency a guarantee for reliable knowledge online? Koivisto discusses two intertwined themes: transparency as an inconspicuous medium, and the ways in which our digital realities become personally customized. Koivisto argues that the way in which transparency discourse participates in digital reality production makes it even more difficult to decipher the “ground truth” beyond the digital representation. In the future, it may become impossible to distinguish representation from simulacra. Because digital transparency builds on a twofold mediation of reality—transparency as a medium itself and algorithms as a medium of digital transparency—the promise of transparency, seeing with one's own eyes, no longer guarantees a privileged access to reality.

Riikka Koulu discusses transparency and technological design from the perspectives of law and science and technology studies. To what extent can transparency be embedded in algorithmic systems through design choices, and is transparency as such desirable? Koulu shows that the entire notion of “transparency by design” is unclear. What exactly is the object of transparency by design: the designed product, or the design process itself? Koulu discusses the normative implications and architectural analogies of design, the limitations of “by design” in producing meaningful control over algorithmic power, and the difficulties of implementing values such as transparency into technological design. Through an examination of “transparency by design and default” in the EU's General Data Protection Regulation, she observes that transparency of design is not at the heart of transparency by design in the data protection framework. Koulu argues that instead of focusing too much on transparency as a design priority, it would be better to approach algorithmic design through the procedural language of access to justice.

Stefan Larsson, Anders Jensen-Urstad, and Fredrik Heintz's article builds on empirical analysis. The research question is the following: To what extent are people aware of the quality and quantity of third-party tracking online? To answer this, the authors analyze tracking in Swedish websites in five sectors: media, retail, banking/insurance, public sector and health. In addition to the tracking detection, the study is conducted by interviewing high-trust and low-trust focus groups. The study shows that tracking is omnipresent online, in particular in retail and media websites. Although people would formally accept the tracking, they remain in fact largely unaware of the magnitude and pervasiveness of the tracking or suffer from consent fatigue and digital resignation. This is in line with the premises of the datafied economy, which depends on the collection, sharing, and commodification of the users' personal data. The basic logic of the datafied economy is hard to resist.

Katja de Vries's article is about transparency, automated decision-making and so-called counterfactuals as means for transparency. De Vries's article thus contributes to the ongoing debate on the black box problem in automated decision-making, as well as transparency's potency to solve the problem. In particular, de Vries concentrates on artificial intelligence and, more specifically, on generative machine learning (in contrast to classificatory machine learning). De Vries argues that so-called counterfactual explanations may, perhaps counterintuitively, do a better job in creating understandability than transparency understood as an explanation of what happened. De Vries argues that it is often more valuable to know what should have happened to receive an alternative result than what actually did happen. However, counterfactual explanations are not unequivocally good or bad. De Vries concludes that counterfactuals could work as a good tool in the transparency toolbox if they are presented in an appropriately constructivist way.

Monika Zalnieriute delves into a phenomenon she calls "transparency washing," which is similar to "ethics washing," namely an attempt to make behavior look acceptable by ostentatiously endorsing ethical standards. As the technology giants exert unprecedented power over us, these companies have also started to use transparency vocabulary to justify their actions. Zalnieriute discusses the ways in which this is done by three of these companies: IBM, Facebook and Google. She shows how pervasive and over-inclusive transparency discourse is in commercial settings. Zalnieriute argues that the use of transparency vocabulary is, in fact, an effort by companies to manipulate their image and to avoid hard regulation. Zalnieriute concludes that this "transparency washing" is connected to a liberal governmentality and a wider trend of "procedural fetishism": by concentrating mostly on processes, wider societal developments may remain unnoticed and unaddressed.

Although all the articles are united by a variety of common concerns, they also show how fragmented transparency discourse actually is. Transparency, moreover, can be criticized from many angles. That said, though intertwined and overlapping, three modalities of scholarly interventions can be identified in this special issue. First, in some articles, the authors' main worry is that there is not enough transparency in the black box society, or that it is not currently doing the job it is supposed to do. As de Vries, Larsson et al., and Grochowski et al. argue in their articles, in many digital contexts, transparency is not

necessarily conceived in the way that would cure the lack of information from which data subjects are suffering. Would we need more transparency, or transparency of another kind? This analysis seeks to make transparency better suited for the circumstance of the datafied society. I call it *the reformative approach* to transparency.

Second, transparency is also political, and it can serve impression management purposes. This is well exemplified by Fenster's and Zalnieriute's articles, which delve into the way in which the providers of transparency benefit from joining that discourse formation of legitimacy. Regardless of how the recipients of transparency experience it, we should also pay attention to those who are preaching it. Therefore, transparency needs to be understood not only as a static condition but as a dynamic relation between its provider and recipient. Transparency is thus a speech act: by naming something transparency, transparency may come into being. This may, however, undermine the legal understanding of transparency. I call this *the performative approach* to transparency.

Lastly, the remaining authors concentrate on the conceptual and socio-material implications of transparency itself, and how it shapes power relations in a datafied society. The authors of these articles derive inspiration from continental philosophy and apply it to legal conundrums. With different accents, Albu and Hansen, Koivisto and Koulu discuss the ways in which power circulates in a datafied society with the help of transparency practices. Hence, the focus is not on what information transparency manages or fails to furnish, but the power structures that underpin it. To what extent is transparency itself part of the problem of the black box society? I call this *the conceptual approach* to transparency.

III. Conclusion

As we can see, despite the all-too-positive connotations of transparency, an opposing trend also exists. Critical transparency studies have emerged and become institutionalized over the last ten years or so.⁵ This multidisciplinary slant of scholarship seeks to find out whether or not transparency really works as promised. Jean Starobinski voices well the suspicion that underpins it:

The metaphor of the lifted veil is the figurative counterpart of a realist theory of knowledge. It is an image employed in a naïvely optimistic manner, which pretends to see the true visage behind the mask, to grasp the "thing in itself": to touch the reality behind the appearance, the substance beyond the accident.⁶

Be it a lifted veil or an opened black box, the assumption is the same: the truth is hidden behind the appearance.

This issue represents critical transparency studies with a legal emphasis. As the articles in this special issue show, it is difficult not to have a conflicted view of transparency in the digital environment. It is problematic in many ways, but at the same time, it is hard

⁵ Ida Koivisto, *Towards Critical Transparency Studies*, 25 *Res Publica* 439 (2019).

⁶ Jean Starobinski, Jean Jacques Rousseau: Transparency and Obstruction 74 (Arthur Goldhammer trans., 1988).

to imagine a viable alternative to it; should it be understandability, access to justice, counterfactual explanations, algoactivism, stricter regulation, or a complete change of vocabulary? Each article shows some troubling sides of the transparency saga. Is transparency a problem, a solution, or a problematic solution to the questions of knowing and understanding? Is it a last bastion of the realist theory of knowledge in the world of post-truth politics and fake news? Whatever the case may be, transparency is surrounded by epistemological anxieties.

It is important to continue the discussion on the concepts and values that channel power and legitimacy in the datafied society. Transparency is one of those concepts. Not only naming problems but also naming solutions includes discursive power. Although transparency is over-inclusive to the extent that it approaches being an empty signifier, it still relies on the power of seeing. In the digital environment, this power has lost some of its privilege as the guarantor of knowing. When photographs, texts and perhaps even crime scenes are transformed into zeros and ones, we are merely left with bafflement. We see, but we do not understand. Yet, we—dual citizens of the analog and the digital worlds—continue to search for answers.